

## 指定画像と類似した画像生成のためのデータセットの構築

T150508 近澤 勇太

指導教員 三好 力 教授

### 1. はじめに

今日、インターネットや SNS の利用者数の増加と発展ともに画像を使用する機会が増えている。気軽に画像の投稿を行えるがために、他人の画像や著作権に守られた画像まで簡単に投稿できてしまう。そこで本研究では、SNS やインターネットでの安易な著作権侵害の予防を目的とした、自由に利用できる画像の提供を行うシステムの提案を行う。本システムでは、自らが SNS やインターネット上で使用したい画像を検索クエリとし、類似点のある画像をインターネット上で機械的に集めることを想定し、使いたい画像と類似する画像を Deep Learning の手法の一つである DCGAN(敵対的生成ネットワーク)で生成する。一枚の画像から自動で似て非なる画像を Deep Learning の学習によって生成されたネットワーク、つまり人工知能が合成することで問題を解決する。本実験では類似する画像を検索するにあたり、画像特徴を用いた手段を用いて、画像生成用のデータセットの構築が可能であるかの検証を行った。

### 2. 実験

ORB、A-KAZE、ヒストグラムの3つ特徴量を用いて、鳥の画像1万枚とそのほかの画像1万枚で構成されたデータベースから、類似度の高い画像の検索を行った。まず、類似度の高い画像を上位5000枚取得し、その結果の中から複数の特徴量に共通する画像を調べる。次に最も検索結果の良かったORB、ヒストグラムの検索手法に対して、画像の取得枚数を100、500、1000、2000、3000、4000、

5000枚と区切り、その中に含まれる共通する画像とその中の鳥の画像の枚数を調べる。

### 3. 実験結果

複数の特徴量の取得画像の中で、共通する画像とその中の鳥の画像枚数の比を総取得枚数ごとに分けて、図1に示す。

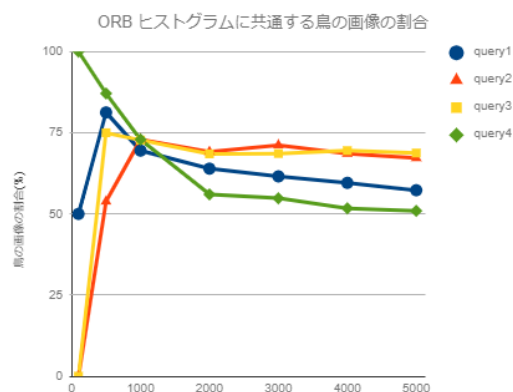


図1 ORB ヒストグラムに共通する鳥の画像の比

### 4. 考察

複数の特徴量を組み合わせることで検索手法では、取得画像5000枚に含まれる鳥の画像の割合は70%を下回る結果となった。少ない取得枚数においては鳥の画像の割合は増えるが共通する画像数自体が非常に少なくなってしまうため、この手法で集めた画像をそのままデータセットとして利用するのは難しいと考える。今後は、実際に学習を行った結果についても検証を行う。